DOI:
10.1128/JCM.01177-18

Document Version
Peer reviewed version

JCM Accepted Manuscript Posted Online 24 October 2018
J. Clin. Microbiol. doi:10.1128/JCM.01177-18

The longitudinal clinical performance of the RNA-based AHPV Human Papillomavirus (HPV) Assay in comparison to the DNA-based Hybrid Capture 2 HPV Test in 2 consecutive screening rounds with a 6-year interval in Germany

Running title: Longitudinal performance of the AHPV test

Thomas Iftner[a]*, Klaus-Joachim Neis[b], Alejandra Castanon[c,d], Rebecca Landy[c], Barbara Holz[a], Astrid Woll-Herrmann[b], Angelika Iftner[a], Annette Staebler[e], Diethelm Wallwiener[f], Claus Hann von Weyhern[e], Felix Neis[f], Juliane Haedicke-Jarboui[a], Peter Martus[g], Sara Brucker[f], Melanie Henes[f], Peter Sasieni[c,d]*

[a] Institute of Medical Virology and Epidemiology of Viral Diseases, University Hospital Tübingen, Tübingen, Germany
[b] Frauenärzte am Staden, Saarbrücken, Germany
[c] Centre for Cancer Prevention, Queen Mary University of London, London, United Kingdom
[d] Cancer Prevention Group, School of Cancer & Pharmaceutical Sciences, Faculty of Life Sciences & Medicine, King's College London
[e] Department of Pathology and Neuropathology, University Hospital Tübingen, Tübingen, Germany
[f] Department of Gynaecology and Obstetrics, University Hospital Tübingen, Tübingen, Germany
[g] Institute for Clinical Epidemiology and Applied Biometry, Medical Faculty Tübingen, Tübingen, Germany.
*corresponding author: Thomas Iftner: thomas.iftner@med.uni-tuebingen.de; Peter Sasieni: peter.sasieni@kcl.ac.uk

Abstract

Longitudinal data on the E6/E7 mRNA-based AHPV® HPV (AHPV) assay exceeding three years in comparison to the gold standard *digene* Hybrid Capture® 2 (HC2) test are not available. We previously reported the cross-sectional data of the German AHPV Screening Trial (GAST) where 10,040 women were recruited and tested by liquid-based cytology, the HC2 and the AHPV assay. 411 test-positive women were followed for up to six years. In addition, 3,295 triple-negative women were screened after a median time of six years. Overall 28 CIN3 cases were detected. The absolute risk of developing high risk HPV positive CIN3+ over six years among those women that tested negative at baseline was 2.2 (1.0-4.9) and 3.1 (1.7-5.7) per 1,000 women screened by the HC2 and the AHPV test, the additional risk in AHPV negative compared with HC2 negative was 0.9 (-0.2 to 2.1) per 1,000, whereas the absolute risk following a negative LBC test was 9.3 (2.9-30.2). The relative sensitivity of AHPV compared to HC2 was 91.5% for CIN3+ and the negative predictive values were 99.8 (99.5-99.9) for HC2 and 99.7 (99.4-99.8) for AHPV.

Our data show that the longitudinal performance of the AHPV-test over six years is comparable to the performance of the HC2 test and that the absolute risk of CIN3+ over six years following a negative AHPV result in a screening population is low.

Keywords

AHPV HPV, cervical cancer screening, E6/E7 mRNA, cervical intraepithelial neoplasia

48 **Introduction**

49 Systematic screening has led to a significant decrease in cervical cancer cases worldwide.
50 Since persistent human papillomavirus (HPV) infection with 13 high risk HPV (HR HPV) types
51 defined as class I or IIA carcinogenic for women (1) is a necessary prerequisite for
52 development of precancerous lesions and cervical neoplasia, tests for HR HPV infections
53 have been developed and validated. Incorporation of molecular HPV testing into cervical
54 cancer screening programmes results in fewer cases of cancer and high-grade cervical
55 intraepithelial neoplasia (CIN3) being detected at the second screening round in those who
56 were tested by HPV-tests compared to women screened by cytology only (2-8). The lower
57 risk following HPV testing suggests that extended screening intervals are appropriate (9-11),
58 which also avoids detection of transient infections in consecutive screening rounds leading
59 to overtreatment.
60 Currently more than 190 HPV assays are commercially available (12) and many countries
61 have implemented HPV tests in their cervical cancer screening programmes, while other
62 countries are in the process of switching from cytological screening to primary HPV testing
63 or co-testing (9, 13, 14), and national cervical cancer screening guidelines have been adapted
64 accordingly (13). Five group tests are approved by the FDA for application in the US and
65 primarily detect the HR HPV-group. To date only the cobas 4800® HPV Test (Roche, USA) and
66 the Onclarity HPV Assay (Becton Dickinson) have been approved for first-line primary
67 screening. The *digene* Hybrid Capture-2 (HC2) high-risk HPV DNA test (QIAGEN®, Hilden,
68 Germany) is considered the gold standard of HPV assays as its performance was validated in
69 a large number of randomized controlled trials and it was the first HPV test receiving FDA
70 approval to screen patients with ASCUS, or in women 30 years and older the HC2 High-Risk
71 HPV DNA Test can be used with Pap to adjunctively screen to assess the presence or absence

of high-risk HPV types. HC2 HPV detection is based on full-length genomic RNA probes for hybridization with the viral DNA of the 13 high-risk HPV types. Longitudinal data showed a very low risk of developing cervical cancer over at least five years in women with a negative HC2 baseline test (15). Four randomized trials have demonstrated that the cumulative incidence of cervical cancer after a median of 6.5 years after a negative HC2 test was lower than the cumulative incidence three years after a normal cytology result (10).

However, cross-hybridization of HC2 with at least 26 additional HPV types of low carcinogenicity or undefined risk, has been detected (16-19), which occasionally can be found even in a CIN3 (20). Epidemiological data suggest that such lesions are unlikely to progress to cervical cancer. This makes the use of HC2 as comparator test according to established consensus guide lines questionable as CIN3 associated with non HR HPV types (and of low progressive potential) will be detected by the HC2, but not by a test with a more stringent analytical specificity that detects predominantly class I/IIa carcinogenic types (21).

As stated by Meijer et al., new HPV tests should demonstrate non-inferior sensitivity and specificity compared with the HC2 test in a representative set of samples from a routine screening population in a cross-sectional setting (22). However, these guidelines specify that they apply to DNA-based tests and whether this guideline could also be applied for validation of RNA-based tests is controversial. A key issue is the definition of a representative sample (of cases of CIN2+) and which parameters are required to allow for the extension of screening intervals. It is argued that RNA positivity is a later event in the natural history of cervical neoplasia than HPV DNA positivity, hence there is a desire to have more longitudinal data regarding CIN2+ incidence following a negative RNA-based test to determine whether it is safe to extend the screening interval after a negative RNA-based test.

Journal of Clinical Microbiology

The AHPV HPV (AHPV) assay (Hologic, SanDiego, USA) is based on target-mediated amplification for detection of viral mRNA. The test detects the mRNA of the two HPV oncogenes E6 and E7 of 14 HPV types, which include all high-risk HPV types targeted by the HC2 assay as well as the class 2B type HPV66 (1, 23). The RNA-based AHPV test has been compared to the DNA-based HPV tests in a number of studies, several are typical screening populations (23, 24). In these studies, the AHPV test consistently demonstrated comparable sensitivities for the detection of CIN2+ or CIN3+, and superior specificity.

We recently reported the cross-sectional results of the German AHPV Screening Trial (GAST), where the clinical sensitivities and specificities of the AHPV and the HC2 HPV tests were determined and compared in cervical samples from 9,451 women aged 30 to 60 years from a routine screening population (24). Samples were centrally analysed by liquid based cytology (LBC), the AHPV assay and the HC2 assay and those women who had a positive result in any of these tests were referred for colposcopy. There was no statistical difference between the AHPV and the HC2 test regarding their sensitivities in detecting CIN2 or CIN3+ lesions. The specificity (<CIN2) and the positive predictive value (CIN2+) of the AHPV test were significantly improved compared to the HC2 test. The GAST study results are in line with a previous report by Heideman et al. which confirmed the non-inferiority of the AHPV assay vs the GP5+/6+ test, and showed that the AHPV test fulfils cross-sectional clinical HPV test requirements for cervical screening (25). Recently, longitudinal clinical performance of the AHPV assay compared to the HC2 test was analysed in a prospective clinical study (CLEAR) including three years of follow-up in 6,201 women (26). Estimated sensitivity of the AHPV test was similar and specificity slightly higher than those of the HC2 test. After three years of

Journal of Clinical Microbiology

follow-up, women, who were HPV-negative (AHPV or HC2) at baseline, had a very low risk of CIN2+ and CIN3+.

However, longitudinal data from a screening population cohort on the AHPV assay exceeding three years compared to the gold standard HC2 testing are important for reassurance especially after the recent introduction of extended screening intervals (≥5 years) in some national cervical cancer screening programs. To address this lack of data, the GAST trial was continued by annually inviting all untreated women who remained positive in at least one of the three tests for follow-up screening. Furthermore, a randomly selected group of 4,000 women who were triple negative at baseline were invited for a second screening round after a mean of six years. We here report the first longitudinal data of more than three years regarding cumulative risk for CIN2/3+, clinical sensitivity and NPV for the detection of histologically reviewed high-grade CIN by the RNA-based AHPV assay in comparison to the HC2 test.

Journal of Clinical Microbiology

Materials and Methods

Participants.

Women aged 30 to 60 years from the routine cervical cancer screening population of three German centres in Tübingen, Saarbruecken, and Freiburg were invited to participate in the GAST trial. The data of the baseline cross-sectional study have previously been published (24). Written informed consent was obtained from each participant, and the study protocol was approved by all relevant ethics committees (Ethik-Kommission Universitätsklinikum Tübingen, reference no. 475/2008MPG1; Ethik-Kommission Alfred Ludwigs-Universität Freiburg, reference no. EK Freiburg 63/09; EthikKommission Landesärztekammer Baden-Württemberg, reference no. B-2009-030f; Ethik-Kommission Ärztekammer des Saarlandes, reference no. 02/10).

Study design.

The design of the baseline cross-sectional study was described previously (24). In brief, eligible consenting women (N = 10,040) had single liquid-based cytology samples (PreservCyt®, Hologic, USA) taken during the annual routine gynaecological examination. Liquid-based cytology (LBC, ThinPrep® Pap Test, Hologic, USA), the *digene* Hybrid Capture 2 (HC2) high-risk HPV DNA test, and the AHPV HPV (AHPV) assay were performed on all samples. All women with a positive result in any of the three screening tests were invited for colposcopy within 8 weeks of receiving their test results.

For the positive follow-up arm of the GAST trial, all women, who tested positive in any of these assays and who were not treated because of abnormal colposcopy and/or histology, were retested annually for up to five years. After a mean interval of six years 4,000 women, who were triple negative at baseline, were randomly selected and invited to be retested by

Journal of Clinical Microbiology

158 all three tests when attending routine cervical screening (i.e. second screening round). Those
159 who tested positive in any of the three tests were invited to colposcopy. Rational for the
160 sample size among women who tested triple negative at baseline can be found in the
161 supplemental methods.
162
163 Liquid-based cytology.
164 As previously described for the cross-sectional trial (24), LBC results were evaluated
165 according to the Munich nomenclature II and translated into the Bethesda System (TBS). LBC
166 results were considered negative when the result was Pap I/II (equivalent to negative for
167 intraepithelial lesion or malignancy [NILM]) or Pap IIw (equivalent to inadequate or atypical
168 cells of undetermined significance [ASCUS]); all other results were considered positive.
169
170 HPV testing.
171 HPV testing was performed as previously detailed (24). Residual LBC samples were
172 processed for HPV testing according to the manufacturer's specifications. Remaining
173 samples were stored for LiPA Extra genotyping in case of positive HPV test results.
174 *digene* Hybrid Capture 2 high-risk HPV DNA testing was performed as described previously
175 (27), using the Rapid Capture® System (RCS, QIAGEN, Hilden, Germany) according to the
176 instructions. A cut-off value of relative light units/cut-off (RLU/CO) ratio of 1.0 for positive
177 test results was used in this study. All PreservCyt® samples with an initial result of >=1 and
178 <2.5 RLU/CO were retested as recommended by the manufacturer. If the retest result was
179 >= 1 RLU/CO, the final result was reported as positive. However, if the retest result was
180 negative, a third test was performed to generate a final two out of three result.

Journal of Clinical Microbiology

The AHPV HPV assay was performed following the manufacturer's instructions. The earlier cut-off value of a signal/cut-off (S/CO) ratio of 1.0 instead of the current (0.5) was used throughout this study to provide continuity of the data. HPV genotyping was carried out using the INNO-LiPA® HPV Genotyping Extra test (Fujirebio, Gent, Belgium), as described previously (27, 28).

Disease Ascertainment and Histopathology.

Women who tested positive in either LBC, the AHPV or the HC2 assay (HPV-positive women) were referred to colposcopy within 8 weeks. If lesions were detected after application of acetic acid a biopsy was taken from the suspicious tissue and specimens were processed to produce H&E stained slides. Current practice in Germany and some other European countries is to observe CIN2 lesions instead of treating them immediately, depending on the individual situation of the patient and her agreement. After local pathologist review, all slides were classified using the three-tiered CIN terminology. All slides with abnormal findings were reviewed by a second pathologist blinded to the first diagnosis and slides with discordant review results were again reviewed by a third pathologist to reach a consensus diagnosis (two out of three agreement).

Statistical analyses.

Prior to analysis, data were plausibility-checked and monitored. This included violation of inclusion criteria (pregnancy, age below 30 years or above 60 years), and positive Pap test six months prior to baseline testing as well as HIV infection. Following this the databases were sealed and sent to the statisticians for statistical analysis.

Journal of Clinical Microbiology

Women with at least one positive test at baseline, who had at least one adequate screening test result on follow-up and who were not treated nor diagnosed with CIN3 or worse at baseline were eligible for the follow-up analysis. In addition, women who tested negative at all three tests at baseline and had at least one adequate screening test result during follow-up were eligible for analysis.

We present baseline demographic characteristics from all participants in the study and for those who attended follow-up. To assess whether there was a statistical difference between groups we used a chi-squared test to compare those attending follow-up with those who were eligible to attend but did not.

Estimating the cumulative risk of CIN3+ (and CIN2+).

The follow-up of women in whom all three baseline screening tests were negative was quite different from that in women who had one or more positive screening test results at baseline. Those with all negative results were rescreened once after approximately 6 years. Women who had a positive result at baseline were invited back at 12-18-month intervals until the results of all three tests were negative or until they were treated for high-grade CIN.

At each visit we estimated the hazard of having CIN3+ (or CIN2+) by multiplying the proportion of tested women who were eligible for colposcopy by the proportion of women attending colposcopy who were diagnosed with CIN3+ (or CIN2+). Further details can be found in the supplemental methods. Having estimated the hazard at each visit, we estimated the cumulative probability of disease after several visits using the Kaplan-Meier (product limit estimator) approach. The variance of the modified Kaplan-Meier estimator was derived

in the same way as the Greenwood formula is derived for the usual Kaplan-Meier estimator. The formula is provided in Supplemental Methods.

We estimated the hazard at baseline separately for each of eight groups based on the result combinations (positive or negative) for each of the three screening test results. Subsequently we estimated the hazards separately in just four groups based on the baseline result combinations of the two HPV tests. There was no evidence that the hazards differed depending on the LBC result within each of the four groups (and numbers were too small to estimate hazards separately in each of the eight groups). Since we didn't have six-year follow-up data for women who were not triple negative at baseline, we assumed that the hazard observed at about 6 years in the baseline screen negative group, also applied to all other groups.

We then estimated the number of cases of CIN3+ (and CIN2+) that would have been observed among women negative at baseline on each of the three tests (separately) had everyone been followed to six years by taking a weighted sum of the estimated cumulative risk in each group. The cumulative risk in each group was also estimated, by dividing the number of cases by the number of women with that result at baseline.

Confidence intervals were obtained by assuming that the logarithm of the cumulative risk is approximately normally distributed. P-values are estimated from the discordant pairs using the exact McNemar significance probability test.

The main analysis presents results including all CIN3 cases. Since we were interested in comparing the performances of two HPV tests both of which aim to detect the same 13 high-risk HPV types and HPV66 and since CIN3 caused by other HPV types are less likely to progress to cancer, a sub-analysis excludes disease where the HC2 results are technically

Journal of Clinical Microbiology

false positive due to cross-hybridization with non-carcinogenic HPV types (i.e., we include only lesions positive for one of the 13 types classified as class I/IIa carcinogenic to humans). Analyses were carried out using STATA 15 (StataCorp, 15.0).

Results

Out of the 4,000 women with negative screening test results at baseline invited, 3,295 (82.4%) attended follow-up. Among those with at least one positive test at baseline 606 were eligible for follow-up and 411 (67.8%) attended (Figure 1). Baseline demographic characteristics of women eligible for analysis at baseline and follow-up are presented in Table 1. Women who participated in the cross-sectional study were broadly similar to those who attended follow-up. There were only slight differences in education and number of sexual partners between women attending follow-up and those eligible, but who did not attend.

Results for women on follow-up after a positive baseline test result.

Untreated women with at least one positive test result at baseline and no CIN3 or worse were invited to attend annual follow-up examinations over a 5 year-period (N=606). Follow-up ceased when HPV infection and/or cervical abnormalities were cleared, if treated for cervical disease, or if they refused to participate in the follow-up study. Of the eligible women 411 (67.8%) attended at least one follow-up examination and were eligible for analysis. The median time to the first follow-up visit was 14 months (range 6 to 80 months) (Figure S1) and the average number of follow-up visits per participant was 1.7 (range 1-5). Three women were excluded, because they were missing a HC2 test and did not return for follow-up. Of the 408 women with at least one follow-up visit with adequate HC2 and AHPV results, 77.2% (315) were negative on both HPV tests at their final visit. In total 200 women tested positive during follow-up and were referred to colposcopy; 165 (82.5%) of these women attended colposcopy. Ninety percent of those who attended follow-up did within 2.5 years of baseline. A total of 32 women were diagnosed with CIN2 or worse during follow-up.

Journal of Clinical Microbiology

Follow-up HPV test results by visit number among those with a positive screening test at baseline are detailed in supplementary Table S1. No LBC test results were missing, but 10 women had both HPV tests missing and 19 were missing the HC2 test on at least one appointment. The agreement of the HPV tests (when both were available) was substantial with a kappa value of κ = 74.7% (95% CI 69.7% to 79.7%).

During follow-up of those women who tested positive on at least one test at baseline, a total of 24 women were diagnosed with CIN3 and 8 with CIN2 (Table 2). Baseline test result and numbers diagnosed with CIN2+ during follow-up are shown in Table 2. 24 CIN2+ lesions (75%) were detected in women with negative cytology and with at least one positive HPV test result and 8 (25%) in women who tested triple positive at baseline. At baseline, HC2 was negative in one CIN2 and two CIN3 cases, whilst AHPV was negative in one CIN2 and five CIN3 cases that developed during follow up (data not shown). One of the 5 CIN3 with discordant HR HC2 positive and AHPV negative HPV test results at baseline was identified by genotyping as HPV 82, which is not targeted by either assay and which is not a HR type. No adenocarcinoma in-situ (AIS) or invasive cervical cancer cases were detected during follow-up.

Longitudinal results for women who tested negative at baseline.

In the baseline cross-sectional arm of the German AHPV Screening Trial (GAST), 8,752 women had a negative result in all three tests (cytological screening, Hybrid Capture 2 (HC2) and AHPV). Of these, 4,000 participants were invited for follow-up testing approximately six years post enrolment. In total 3,295 (82.4%) attended follow-up (Figure 1). The median time between baseline and attendance at the second round was 6.2 years (range 3.9-8.5).

301 At the second round 3,057 women tested negative on all three tests (92.8%). A total of 140
302 women (4.6%) had at least one positive test results at follow-up, 115 (82%) of these
303 underwent a colposcopic examination and a total of 9 women were diagnosed with CIN2 or
304 worse disease (5 CIN2 and 4 CIN3 lesions). A summary of LBC and HPV tests results at the
305 second screening round is found in Table 3. The level of agreement between the HPV tests
306 was substantial with a kappa value of κ = 81.1 (95%CI: 78.0-93.8).
307 Sensitivity of cytology for the detection of CIN3+ was 44% (N=4 of 9), but 100% tested HPV
308 positive (Table 4). One CIN3 case, which tested HC2 positive and AHPV negative, revealed in
309 the histopathology a small lesion of 0.2 mm that was regressive and showed signs of
310 inflammation. HPV 16 was detected in all patients with CIN3 by LiPA- Extra genotyping test.
311 In the present study we observed 10 of 23 untreated (43%) CIN2 cases that regressed, while
312 3/23 (13%) progressed to CIN3.
313 Passive clinical follow-up data were available from a registry on the complete Saarbrücken
314 sub-cohort of 2,147 women who tested triple-negative at baseline, 887 of those women
315 attended follow-up as part of GAST. During a six years passive follow-up period only one
316 CIN1 and one CIN2 case were observed in women who did not attend the second-round
317 screening in GAST. Among the Saarbrücken cohort attending the second screening round,
318 one case of CIN2 and two cases of CIN3 were detected at the second screening appointment.
319
320 **HPV-types in samples with high grade disease.**
321 All HPV-positive samples were genotyped by the LiPA-Extra genotyping test. Baseline HPV
322 test results among the 41 women who went on to be diagnosed with CIN2+ during follow-up
323 show that 9 (22%) were HPV negative, 2 (5%) tested positive to non-HC2 risk HPV types (66

and 82), 24 (58%) single and 8 (20%) multiple HPV infections were detected (results not shown).

HPV genotyping results at the time of diagnosis (during follow-up) are presented in Table S2. At the time of diagnosis 1/41 (2%) CIN2 was HPV negative on both tests, one CIN2 and one CIN3 (5%) tested positive to non-high risk HPV types (53, 66 and 82), 28/41 (68%) single and 12/41 (29%) multiple HPV infections were detected.

HPV16 was the most frequent HPV-type detected in patients with CIN3 in the cross-sectional part of the study and among those attending follow-up.

Cumulative risk of disease during the study period.

The main analysis presents results including all diagnosed disease. We present a sub-analysis excluding two CIN3 cases whose HPV types were 82 and 67 and hence were considered technically false positive HC2 test results. One case (HPV 67) was diagnosed and treated at baseline, the remaining case (HPV 82) was diagnosed at follow-up.

A summary of the 6-year cumulative risk per 1,000 women screened and negative predictive value among women testing negative at baseline can be found in Table 5. Risk per 1,000 women screened by time since baseline test is presented in Figure 2 and 3. Note that the vast majority of women negative on any one screening test were negative on all three and were therefore not rescreened until 6 years. This explains the sudden jump in the risk at 6-year visits.

CIN2 or worse.

Cumulative risk of CIN2 or worse by the 6-year visit was 0.62% (95%CI: 0.24% to 1.59%) and 0.47% (95%CI: 0.27% to 0.81%) among those who tested AHPV and HC2 negative, respectively. The difference in AHPV negative was 0.15% (95%CI: 0.38% less to 0.69% more) and is not significant (p= 0.096). For comparison the cumulative risk by 6 years among LBC negative women was 1.66% (0.72% to 3.83%). The relative sensitivity for CIN2+ of AHPV in comparison to HC2 was 91.4%. Among women testing negative on both HPV tests at baseline, the cumulative risk of CIN2 or worse was 0.38% (95%CI: 0.17% to 0.86%).

The sub-analysis excluding one case (diagnosed at follow-up) of CIN3 which tested HPV 82 positive and one (diagnosed at baseline) that tested HPV 67 positive, produced very similar results : 0.59% (0.22% to 1.61%) and 0.47% (0.27% to 0.81%) among AHPV and HC2 negative women, respectively. The relative sensitivity for CIN2+ of AHPV in comparison to HC2 was 93.0%.

CIN3.

Cumulative risk of CIN3 disease by the year-6 visit was 0.31% (95%CI: 0.17% to 0.57%) and 0.22% (95%CI: 0.10% to 0.49%) for AHPV negative and HC2 negative women (Table 5), respectively: difference 0.09% (95%CI -0.02% to 0.21%). The cumulative risk by the year-6 visit among those testing LBC negative was 0.93% (0.29% to 3.02%). The relative sensitivity for CIN3 of the AHPV test in comparison to HC2 was 91.5%. Among women testing negative on both HPV tests at baseline, the cumulative risk of CIN3 was 0.17% (95%CI: 0.04% to 0.75%).

The sub-analysis excluding two CIN3 cases with technically false positive HR HC2 HPV type results, produced cumulative risks by 6 years of 0.28% (0.14% to 0.54%) among those who tested AHPV negative at baseline and 0.22% (0.10% to 0.49%) among those who tested HC2 negative (p= 0.1094). The cumulative risk by 6 years among those testing LBC negative was 0.90% (0.27% to 3.04%). The relative sensitivity of AHPV to HC2 for CIN3 increased to 94.2%. There were only 20 women with a signal/cut-off ratio of between 0.5 and 1.0 on the AHPV test at baseline, they were all HC2 negative and LBC negative. Only four of these 20 women attended follow-up where they were found to still be HPV negative.

Journal of Clinical Microbiology
JCM
Journal of Clinical Microbiology

Discussion

In recent years, many countries integrated HPV testing into their national cervical cancer screening programmes. Compared to conventional methods, HPV testing increases early detection rates of precancerous and cancerous lesions and allows extended screening intervals. However, the optimal lengths of screening intervals for women with negative results remains to be established and might greatly depend on the long-term predictive values of a given HPV test. Longitudinal clinical performance data have so far been published for only a small number of HPV tests. Ronco et al. presented pooled data from four studies on the performance of the DNA-based HC2 assay over a median of 6.5 years follow-up period (10). In addition, there is evidence regarding the good negative predictive value over three years for the cobas 4800® test (Roche Diagnostics) (29), over three years for the Abbott RealTime HPV DNA-test (30) and over three years for the RNA-based AHPV test (26). During the revision of our manuscript, data comparing the AHPV with the cobas 4800 HPV test using biobanked material were published that demonstrate a non-inferior longitudinal sensitivity and NPV over 7 years for the AHPV (31).

In the present study we evaluated the extended predictive value of the RNA-based AHPV HPV test in comparison to the DNA-based HC2 test over a 5-6-year period by focussing on the cumulative risks for CIN3+ six years after a negative baseline result. In our opinion CIN2+ is a less reliable endpoint because it is an equivocal histological diagnosis and regression rates are high, as observed in our study with a percentage of 43%. An advantage of this study therefore was that many CIN2 lesions were not treated immediately and were seen to regress during surveillance.

During the course of the follow-up of women who tested positive (LBC, AHPV or HC2) at baseline, we detected 8 CIN2 and 24 CIN3 cases. One CIN2 case was missed by both HPV

Journal of Clinical Microbiology

tests and was positive only by cytology. One CIN3 case tested negative by AHPV at the time of diagnosis, but was detected by HC2 and it contained a non-HR type (HPV82). Results from a meta-analysis of type specific HPV DNA prevalence in cervical cancer and women with normal cytology showed a prevalence of HPV82 of 0.1% (95% CI 0.1-0.3) and 0.1% (95% CI 0.0-0.1), respectively. HPV82 is not targeted by the HC2 test, but may yield positive results due to cross-reaction. The known extensive cross-reactivity of the HC2 test may therefore explain the non-significantly higher sensitivity compared to the AHPV test in the baseline and follow up results of this study. According to the Meijer criteria (22) the candidate test should have a clinical sensitivity for CIN2+ not lower than 90% of the clinical sensitivity of the HC2 in women aged at least 30 years. Clearly the results for AHPV at 6 years achieved clinical sensitivity rates exceeding this 90% threshold, regardless if all CIN2+ cases were included or if CIN2+ with non-carcinogenic types were excluded.

In the second screening round of women who tested triple negative at baseline a total of five CIN2 and four CIN3 cases were identified of which one CIN3 was missed by the AHPV test at follow-up. This case was HPV16-positive and was detected repeatedly positive by the HC2 test at relative light units (RLUs) of 1.86, 1.53 and 1.51, which is a borderline positive result according to the FDA approval, but a negative test result in some countries (e.g. United Kingdom), where an increased cut-off of 2.0 RLUs is used for cervical cancer screening.

The cumulative risks of CIN3+ 6 years after a negative screening test in this study are very similar to the ones observed earlier by Dillner et al. (15). In both studies, the cumulative incidence in women with baseline negative cytology was around 1%. The cumulative incidence after a negative AHPV in this study was substantially lower: 0.31% (0.17% to 0.57%). This is in line with one previous publication where very low three-year risks for

Journal of Clinical Microbiology

CIN3+ were detected after a negative baseline AHPV test (26). In our study, the upper 95% confidence limit for the additional risk after a negative AHPV compared with a negative HC2 was 0.21%. If it is accepted that women do not need to be re-screened until their risk of CIN3+ reaches 0.5%, this study has shown that it is safe to use an interval of 5-6 years after a negative AHPV test.

The analyses including all CIN3 cases detected or excluding two CIN3 (one from baseline with HPV67 and one detected at follow up with HPV82) revealed highly comparable absolute risks for CIN3+ following a negative baseline HC2 or AHPV test and comparable longitudinal negative predictive values (NPV) of both tests. Note that women who tested positive (on any test) at baseline were not asked to return for testing 6 years after enrolment. Therefore we assumed that the hazard at 6 years among those who tested positive but who did not develop disease was the same as that observed at about 6 years in the baseline screen negative group. This explains the sudden jump in the risk in figure 2 and 3 after 5 years.

Altogether, we found an absolute risk for developing CIN3+ after six years among those women who tested negative at baseline of 2.2 and 3.1 per 1,000 women screened by the HC2 and the AHPV test, respectively. This difference is not significant and is in line with one previous publication where very low three-year risks for CIN3+ were detected after a negative baseline AHPV test (26).

The fact that annual follow-up of all women from the routine screening population was unachievable complicated the data analysis and might be considered a weakness of this study. However, the CLEAR study found that annual follow up of screen-negative women suffers from low compliance (26) and is not recommended by national and European guide lines. Passive follow-up of those women who tested triple-negative at baseline was possible in a subset of 2,147/8,752 (25%) women and showed that we were unlikely to have missed

Journal of Clinical Microbiology

disease by only rescreening at 6 years. On the other hand, this study strongly benefits from the large number of participating women, the prospective study design and the extended follow-up period of six years. Another strength of this study is that all positive samples (LBC, AHPV or HC2) were genotyped, which enabled a detailed analysis of discordant test results. The poor sensitivity of LBC in this study may be considered a weakness, but simply reflects the low single-round sensitivity of cytology in Germany that has been noted in several studies previously (32). Since the primary comparison is between AHPV and HC2 the poor sensitivity of LBC does not affect the overall results.

In summary, numerous studies from different populations (23) consistently demonstrated a similar cross-sectional sensitivity paired with higher clinical specificity when AHPV was compared to other FDA approved HPV DNA tests, which reduces the costs of follow-up. With regard to the extended intervals in some cervical cancer screening programs data for screening intervals up to 3 years have already been published, as well as a retrospective analysis over 7 years (31). With the present study we add prospective data of the longitudinal performance over a 5-6 year period showing that the cumulative risk for CIN2/3 and the NPV of the AHPV is non-significantly different from the HC2 assay. We conclusively demonstrate low absolute risks for CIN3+ following a negative AHPV test suggesting that the extended screening intervals proposed for use with HC2 are safe with AHPV too.

Journal of Clinical Microbiology

Acknowledgements

This study was supported by an unconditional grant from Hologic Inc. to the University Hospital Tübingen. Hologic had neither influence on the study protocol nor access to the primary data and was not involved in the analysis and writing of the manuscript. AC and RL were supported by Cancer Research UK grant A16892 to PS.

Finally, we thank all collaborating gynaecologists, as well as all women who participated in this study.

Registration

This trial is registered at clinicaltrials.gov, Identifier NCT02634190.

References

1. IARC. 2012. A Review of Human Carcinogens - Biological Agents, vol 100B.
2. Bulkmans NW, Berkhof J, Rozendaal L, van Kemenade FJ, Boeke AJ, Bulk S, Voorhorst FJ, Verheijen RH, van Groningen K, Boon ME, Ruitinga W, van Ballegooijen M, Snijders PJ, Meijer CJ. 2007. Human papillomavirus DNA testing for the detection of cervical intraepithelial neoplasia grade 3 and cancer: 5-year follow-up of a randomised controlled implementation trial. Lancet 370:1764-72.
3. Kitchener HC, Almonte M, Thomson C, Wheeler P, Sargent A, Stoykova B, Gilham C, Baysson H, Roberts C, Dowie R, Desai M, Mather J, Bailey A, Turner A, Moss S, Peto J. 2009. HPV testing in combination with liquid-based cytology in primary cervical screening (ARTISTIC): a randomised controlled trial. Lancet Oncol 10:672-82.
4. Naucler P, Ryd W, Tornberg S, Strand A, Wadell G, Elfgren K, Radberg T, Strander B, Johansson B, Forslund O, Hansson BG, Rylander E, Dillner J. 2007. Human papillomavirus and Papanicolaou tests to screen for cervical cancer. N Engl J Med 357:1589-97.
5. Ronco G, Giorgi-Rossi P, Carozzi F, Confortini M, Dalla Palma P, Del Mistro A, Ghiringhello B, Girlando S, Gillio-Tos A, De Marco L, Naldoni C, Pierotti P, Rizzolo R, Schincaglia P, Zorzi M, Zappa M, Segnan N, Cuzick J, New Technologies for Cervical Cancer screening Working G. 2010. Efficacy of human papillomavirus testing for the detection of invasive cervical cancers and cervical intraepithelial neoplasia: a randomised controlled trial. Lancet Oncol 11:249-57.
6. Ronco G, Giorgi-Rossi P, Carozzi F, Dalla Palma P, Del Mistro A, De Marco L, De Lillo M, Naldoni C, Pierotti P, Rizzolo R, Segnan N, Schincaglia P, Zorzi M, Confortini M, Cuzick J, New Technologies for Cervical Cancer screening Working G. 2006. Human papillomavirus testing and liquid-based cytology in primary screening of women younger than 35 years: results at recruitment for a randomised controlled trial. Lancet Oncol 7:547-55.
7. Castle PE, Stoler MH, Wright TC, Jr., Sharma A, Wright TL, Behrens CM. 2011. Performance of carcinogenic human papillomavirus (HPV) testing and HPV16 or HPV18 genotyping for cervical cancer screening of women aged 25 years and older: a subanalysis of the ATHENA study. Lancet Oncol 12:880-90.
8. Stoler MH, Wright TC, Jr., Sharma A, Apple R, Gutekunst K, Wright TL, Group AHS. 2011. High-risk human papillomavirus testing in women with ASC-US cytology: results from the ATHENA HPV study. Am J Clin Pathol 135:468-75.
9. Arbyn M, Ronco G, Anttila A, Meijer CJ, Poljak M, Ogilvie G, Koliopoulos G, Naucler P, Sankaranarayanan R, Peto J. 2012. Evidence regarding human papillomavirus testing in secondary prevention of cervical cancer. Vaccine 30 Suppl 5:F88-99.
10. Ronco G, Dillner J, Elfstrom KM, Tunesi S, Snijders PJ, Arbyn M, Kitchener H, Segnan N, Gilham C, Giorgi-Rossi P, Berkhof J, Peto J, Meijer CJ, International HPVswg. 2014. Efficacy of HPV-based screening for prevention of invasive cervical cancer: follow-up of four European randomised controlled trials. Lancet 383:524-32.
11. Sasieni P, Cuzick J. 2002. Could HPV testing become the sole primary cervical screening test? J Med Screen 9:49-51.
12. Poljak M, Kocjan BJ, Ostrbenk A, Seme K. 2016. Commercially available molecular tests for human papillomaviruses (HPV): 2015 update. J Clin Virol 76 Suppl 1:S3-S13.
13. Saslow D, Solomon D, Lawson HW, Killackey M, Kulasingam SL, Cain J, Garcia FA, Moriarty AT, Waxman AG, Wilbur DC, Wentzensen N, Downs LS, Jr., Spitzer M, Moscicki AB, Franco EL, Stoler MH, Schiffman M, Castle PE, Myers ER, American Cancer S, American Society for C, Cervical P, American Society for Clinical P. 2012. American Cancer Society, American Society for Colposcopy and Cervical Pathology, and American Society for Clinical Pathology screening guidelines for the prevention and early detection of cervical cancer. Am J Clin Pathol 137:516-42.

Journal of Clinical Microbiology

14. Castle PE, de Sanjose S, Qiao YL, Belinson JL, Lazcano-Ponce E, Kinney W. 2012. Introduction of human papillomavirus DNA screening in the world: 15 years of experience. Vaccine 30 Suppl 5:F117-22.
15. Dillner J, Rebolj M, Birembaut P, Petry KU, Szarewski A, Munk C, de Sanjose S, Naucler P, Lloveras B, Kjaer S, Cuzick J, van Ballegooijen M, Clavel C, Iftner T, Joint European Cohort S. 2008. Long term predictive values of cytology and human papillomavirus testing in cervical cancer screening: joint European cohort study. BMJ 337:a1754.
16. Castle PE, Solomon D, Wheeler CM, Gravitt PE, Wacholder S, Schiffman M. 2008. Human papillomavirus genotype specificity of hybrid capture 2. J Clin Microbiol 46:2595-604.
17. Iftner T, Villa LL. 2003. Chapter 12: Human papillomavirus technologies. J Natl Cancer Inst Monogr:80-8.
18. Castle PE, Schiffman M, Burk RD, Wacholder S, Hildesheim A, Herrero R, Bratti MC, Sherman ME, Lorincz A. 2002. Restricted cross-reactivity of hybrid capture 2 with nononcogenic human papillomavirus types. Cancer Epidemiol Biomarkers Prev 11:1394-9.
19. Poljak M, Marin IJ, Seme K, Vince A. 2002. Hybrid Capture II HPV Test detects at least 15 human papillomavirus genotypes not included in its current high-risk probe cocktail. J Clin Virol 25 Suppl 3:S89-97.
20. Bzhalava D, Guan P, Franceschi S, Dillner J, Clifford G. 2013. A systematic review of the prevalence of mucosal and cutaneous human papillomavirus types. Virology 445:224-31.
21. Preisler S, Rebolj M, Ejegod DM, Lynge E, Rygaard C, Bonde J. 2016. Cross-reactivity profiles of hybrid capture II, cobas, and APTIMA human papillomavirus assays: split-sample study. BMC Cancer 16:510.
22. Meijer CJ, Berkhof J, Castle PE, Hesselink AT, Franco EL, Ronco G, Arbyn M, Bosch FX, Cuzick J, Dillner J, Heideman DA, Snijders PJ. 2009. Guidelines for human papillomavirus DNA test requirements for primary cervical cancer screening in women 30 years and older. Int J Cancer 124:516-20.
23. Haedicke J, Iftner T. 2016. A review of the clinical performance of the Aptima HPV assay. J Clin Virol 76 Suppl 1:S40-S48.
24. Iftner T, Becker S, Neis KJ, Castanon A, Iftner A, Holz B, Staebler A, Henes M, Rall K, Haedicke J, von Weyhern CH, Clad A, Brucker S, Sasieni P. 2015. Head-to-Head Comparison of the RNA-Based Aptima Human Papillomavirus (HPV) Assay and the DNA-Based Hybrid Capture 2 HPV Test in a Routine Screening Population of Women Aged 30 to 60 Years in Germany. J Clin Microbiol 53:2509-16.
25. Heideman DA, Hesselink AT, van Kemenade FJ, Iftner T, Berkhof J, Topal F, Agard D, Meijer CJ, Snijders PJ. 2013. The Aptima HPV assay fulfills the cross-sectional clinical and reproducibility criteria of international guidelines for human papillomavirus test requirements for cervical screening. J Clin Microbiol 51:3653-7.
26. Reid JL, Wright TC, Jr., Stoler MH, Cuzick J, Castle PE, Dockter J, Getman D, Giachetti C. 2015. Human papillomavirus oncogenic mRNA testing for cervical cancer screening: baseline and longitudinal results from the CLEAR study. Am J Clin Pathol 144:473-83.
27. Boehmer G, Wang L, Iftner A, Holz B, Haedicke J, von Wasielewski R, Martus P, Iftner T. 2014. A population-based observational study comparing Cervista and Hybrid Capture 2 methods: improved relative specificity of the Cervista assay by increasing its cut-off. BMC Infect Dis 14:674.
28. Petry KU, Luyten A, Justus A, Iftner A, Strehlke S, Schulze-Rath R, Iftner T. 2012. Prevalence of low-risk HPV types and genital warts in women born 1988/89 or 1983/84 -results of WOLVES, a population-based epidemiological study in Wolfsburg, Germany. BMC Infect Dis 12:367.
29. Wright TC, Stoler MH, Behrens CM, Sharma A, Zhang G, Wright TL. 2015. Primary cervical cancer screening with human papillomavirus: end of study results from the ATHENA study using HPV as the first-line screening test. Gynecol Oncol 136:189-97.

30. Poljak M, Ostrbenk A, Seme K, Sterbenc A, Jancar N, Vrtacnik Bokal E. 2016. Three-year longitudinal data on the clinical performance of the Abbott RealTime High Risk HPV test in a cervical cancer screening setting. J Clin Virol 76 Suppl 1:S29-S39.
31. Forslund O, Miriam Elfstrom K, Lamin H, Dillner J. 2018. HPV-mRNA AND HPV-DNA DETECTION IN SAMPLES TAKEN UP TO SEVEN YEARS BEFORE SEVERE DYSPLASIA OF CERVIX UTERI. Int J Cancer doi:10.1002/ijc.31819.
32. McCrory DC, Matchar DB, Bastian L, Datta S, Hasselblad V, Hickey J, Myers E, Nanda K. 1999. Evaluation of cervical cytology. Evid Rep Technol Assess (Summ):1-6.

592 **Figure legends**

593 Figure 1: Follow-up Flow chart. Green indicates the cohort which tested triple-negative (LBC,
594 HC2 and AHPV) at baseline. Red indicates the cohort who tested positive in at least one test
595 (LBC, HC2 or AHPV) at baseline.

596

597 Figure 2: Cumulative risk per 10,000 women becoming CIN2+ following a negative baseline
598 result in the respective test. Data have been analysed by visit. Visits should have been
599 annual up to 5 years in those with a positive test and at 6 years for triple negatives at
600 baseline. LBC, liquid based cytology.

601

602 Figure 3: Cumulative risk per 10,000 women becoming CIN3+ following a negative baseline
603 result in the respective test. Data have been analysed by visit. Visits should have been
604 annual up to 5 years in those with a positive test and at 6 years for triple negatives at
605 baseline. LBC, liquid based cytology.

606

Journal of Clinical Microbiology
JCM
Journal of Clinical Microbiology

607 **Tables**

Table 1. Baseline demographic characteristics of women in the GAST Trial

Characteristic reported at baseline	Attended follow-up		Eligible for analysis at baseline		Chi2 test*
Age at enrolment	**N**	**%**	**N**	**%**	
30-34	611	16.5	1623	17.2	
35-39	692	18.7	1696	17.9	
40-44	852	23.0	2123	22.5	
45-49	734	19.8	1873	19.8	
50-54	499	13.5	1295	13.7	
55-59	318	8.6	841	8.9	
Missing	0	-	0	-	
Total (Not missing)	**3706**		**9451**		**x^2_5=5.163, p=0.396**
Education					
Missing	384	-	1663	-	
None	13	0.4	34	0.4	
Primary	602	18.1	1383	17.8	
College	1503	45.2	3350	43.0	
University	1204	36.2	3021	38.8	
Total (Not missing)	**3322**		**7788**		**x^2_3=17.16, p=0.001**
Number of sexual partners					
Missing	1016	-	3107	-	
One	921	34.2	2195	34.6	
Two to four	1019	37.9	2282	36.0	
Four or more	750	27.9	1867	29.4	
Total (Not missing)	**2690**		**6344**		**x^2_2=8.6423, p=0.013**
Age at first sexual intercourse					
Missing	737	-	2448	-	
Under age 18	1660	55.9	3855	55.0	
Age 18 or older	1310	44.1	3148	45.0	
Total (Not missing)	**2970**		**7003**		**x^2_1=1.4558, p=0.228**

*Note the X^2 test compares those attending to those not attending among those eligible for follow-up

608

609

Table 2. Baseline HPV test results among women with CIN2+ during follow-up

LBC	HC2	AHPV	CIN2	CIN3
+	+	+	2	6
-	+	+	4	11
-	+	-	1	5*
-	-	+	1	2

*One case was positive for HPV82 on LIPA and is excluded in the sub- analysis

610

611

Table 3. Second round LBC and HPV test results among women who were triple negative at baseline

HPV test result during follow-up (HC2/AHPV)	LBC negative	LBC inadequate	LBC low-grade (Pap III)	LBC high-grade (Pap IIID)	Total	Number with CIN2+
Both missing	71	4	0	0	75	0
Missing HC2	1	0	0	0	1	0
Missing AHPV	4	0	0	0	4	0
-/-	3057	18	5	12	3092	0
-/+	13	0	0	0	13	0
+/-	48	0	1	1	50	1
+/+	44	0	3	13	60	8
Total	3238	22	9	26	3295	9

612

613

Table 4. Second round screening HPV test results among women with CIN2+ during follow-up

LBC	HC2	AHPV	CIN2	CIN3+
+	+	+	3	1
-	+	+	2	2
-	+	-	0	1
-	-	+	0	0

Note the one CIN3+ with discordant HPV test results was also negative on LBC as were two other CIN3 and two CIN2.

614

615

Journal of Clinical Microbiology

JCM

Journal of Clinical Microbiology

616

Table 5. 6-year cumulative incidence, risk per 1000 women screened and negative predictive value among those testing negative at baseline.

	Cumulative incidence, 95% CI				Risk per 1,000 women screened, 95% CI				Negative predictive value*, 95% CI			
CIN2 or worse												
AHPV Negative	0.62%	0.24%	to	1.59%	6.2	2.4	to	15.9	99.38%	98.41%	to	99.76%
HC2 Negative	0.47%	0.27%	to	0.81%	4.7	2.7	to	8.1	99.53%	99.19%	to	99.73%
LBC Negative	1.66%	0.72%	to	3.83%	16.6	7.2	to	38.3	98.34%	96.17%	to	99.28%
CIN3 or worse												
AHPV Negative	0.31%	0.17%	to	0.57%	3.1	1.7	to	5.7	99.69%	99.43%	to	99.83%
HC2 Negative	0.22%	0.10%	to	0.49%	2.2	1.0	to	4.9	99.78%	99.51%	to	99.90%
LBC Negative	0.93%	0.29%	to	3.02%	9.3	2.9	to	30.2	99.07%	96.98%	to	99.71%

*Note the NPV is estimated excluding the risk among those attending the second round of screening

617